MAXIMUM-LIKELIHOOD-BASED EXTENDED-IMAGE SPATIAL ACQUISITION AND TRACKING TECHNIQUE*

Haiping Tsou, Caroline Racho and Tsun-Yee Yan

Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109

Abstract – This paper describes a generic extended-image spatial acquisition and tracking technique developed to enable highly accurate and stable pointing to a moving target through an imaging device. It is intended to be a fine-pointing scheme complimentary to the imager's general control subsystem that is able to provide a "coarse" pointing capability such that the target remains within the imager's field of view. This correlation-type scheme compares the received image obtained from the detector array with a priorly established reference in the transform domain to estimate the target's movement. The received image is assumed to be an extended-image covering more than one element of the detector array and have each of its pixels corrupted by an independent additive white Gaussian noise. The coordinate of the target area is acquired and tracked, respectively, by an open-loop acquisition algorithm and a closed-loop tracking algorithm derived from the maximum likelihood criterion. The resulting movement estimate is used to direct the imager to closely follow the moving target. This technique has many potential applications, including free-space optical communications and astronomy where accurate and stabilized optical pointing is essential.

INTRODUCTION

Precision optical pointing is an essential capability of any laser-based system, especially when a laser signal is delivered to or received from a distant moving target. The pointing error caused by factors such as uncompensated platform motion and/or jitter, sensor noise and/or bias, and atmospheric propagation effects, such as image dancing, blurring, and scintillation [1, 2], inevitably impairs the efficiency and, therefore, causes significant system performance degradation. For example, an optical link in deep-space laser communications may extend as long as several astronomy units[1], requiring high-power lasers with very narrow beam divergence for high-rate data transmission. It is not uncommon that the required pointing accuracy in this case is on the order of microradians on both ends of the optical link to ensure the communication quality. Such a stringent requirement posts a real challenge in design of the pointing control system, especially for the optical transceiver onboard spacecraft.

A typical pointing control for optical transceivers used in deep-space applications is designed as a two-level scheme. The coarse-pointing control is capable of keeping the target within the field of view of the optical detector by maintaining the telescope's pointing direction according to some pre-determined parameters such as the predicted target trajectory, etc. On the other hand, the fine-pointing control is intended to track out any residual pointing error not being removed by the coarse-pointing. Normally, it is accomplished by driving a two-axis fast-steering mirror through an extended-image acquisition and tracking such that the target image remains "fixed" on the focal plane detector array. Figure 1 shows the functional block diagram of an optical transceiver for laser communications.

In this paper, a generic extended-image spatial ac-

[1]One astronomy unit is roughly 149.6×10^6 kilometers.

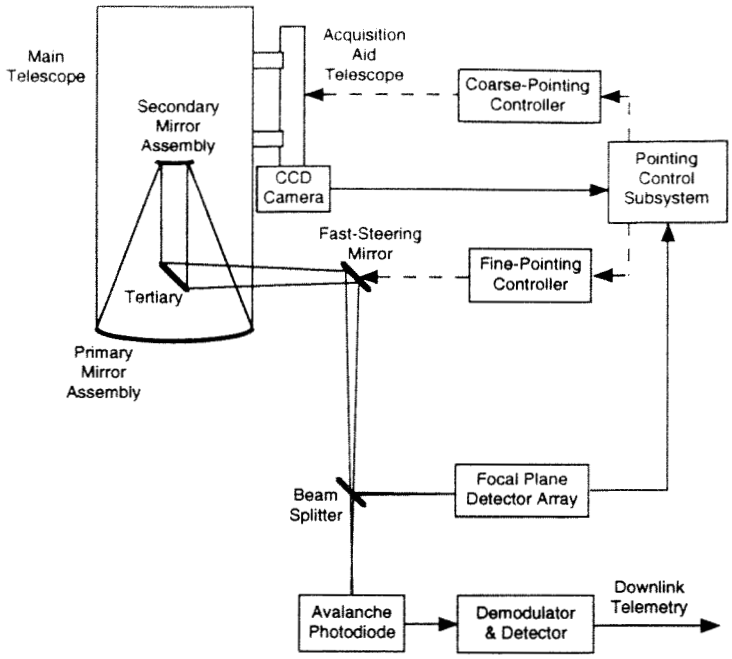


Figure 1: The optical transceiver for laser communications.

quisition and tracking technique developed to enable highly accurate and stable pointing to a moving target through an imaging device is presented. This work has been inspired by studies known for the image registration and the moving target indication [3, 4, 5] and, in some aspect, can be seen as an extension beyond those work by adding a random disturbance, which is modeled as an independent additive white Gaussian noise, to each pixel of the received image. The resulting image is, therefore, a randomly disturbed profile of a known target image which covers more than one element of the imager's detector array. The proposed scheme estimates and tracks the offset of the target's coordinate based on the maximum likelihood criterion derived from the transform-domain correlation between the received image and a priorly established reference profile. It is assumed that the reference profile of the moving target exists or, at least, is able to be synthesized based on the current target position by using target information pre-stored and/or accumulated through continuous tracking. This assumption is generally valid when the relative motion between the target and the imager is largely confined on a two-dimensional plane during the acquisition and tracking[2].

[2]Note that a three-dimensional relative motion viewed by the imager can be treated as on a two-dimensional plane as long as the distance between target and imager is much larger than the range of distance variation between them.

For rotation-invariant relative motions, the optimal acquisition requires solving two simultaneous nonlinear equations to estimate the coordinate offset. A suboptimal estimate exists by solving linearized version of the maximum likelihood criterion when computation complexity needs to be reduced. A closed-loop tracking algorithm motivated by the maximum likelihood criterion is developed as well for continuously tracking the translation movement between the target and the imager. The loop feedback signals of this image tracking loop are formulated as weighted transform-domain correlations between the received image and the previously estimated reference image. Only two simultaneous linear equations are involved in each iteration of the continuous tracking. When a rotational movement is also involved, a maximum-likelihood estimate of rotation angle has to be performed after translation estimation or each iteration of translation tracking. Such a one-shot estimation requires a parallel search through the range of all possible rotation angles to determine the one which is most likely.

In this paper, we will describe the analytical work in details, including a general description of the mathematical model for image representation, the effects of a translational and rotational movement in the discrete Fourier transform domain, the maximum likelihood estimator for spatial acquisition, and the image tracking loop developed for spatial tracking.

The numerical results of sample scenarios to acquire a Sun-lit Earth beacon and to track a laser spot image modeled as a Gaussian pulse are also included, which demonstrate the capability to achieve subpixel resolutions by this maximum-likelihood based spatial acquisition and tracking scheme in a high disturbance environment.

Mathematical Background

Representation of an Extended Image

The received image detected by an $M \times N$ focal plane array at time t_l, denoted as $r_l(m,n)$, can be represented by a sum of the source image, $s_l(m,n)$, and the random disturbance, $n_l(m,n)$, as follows

$$r_l(m,n) = s_l(m,n) + n_l(m,n), \tag{1}$$

where $m = 0, 1, \ldots, M-1$ and $n = 0, 1, \ldots, N-1$. With an additive white Gaussian random disturbance model, $n_l(m,n)$ is assumed to be an independent zero-mean Gaussian random variable with variance σ_l^2 for all m and n.

The discrete Fourier transform of the received image at time t_l becomes

$$\mathcal{R}_l(m,n) = \mathcal{S}_l(m,n) + \mathcal{N}_l(m,n) \tag{2}$$

where the transform-domain source image and random disturbance are

$$\mathcal{S}_l(m,n) = \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} s_l(p,q) e^{-i2\pi\left(\frac{m}{M}p + \frac{n}{N}q\right)} \tag{3}$$

and

$$\mathcal{N}_l(m,n) = \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} n_l(p,q) e^{-i2\pi\left(\frac{m}{M}p + \frac{n}{N}q\right)} \tag{4}$$

Effect of Image Movement on the Focal Plane

When, between t_l and t_{l+1}, the received image translates by the amount of x_l and y_l pixels along the x-axis and y-axis and rotates by an angle δ_l on the focal plane detector array, the resulting image at t_{l+1} is related to the previous image at t_l by

$$s_{l+1}(m,n) = s_l(m',n') + \epsilon_T(m',n') \tag{5}$$

where

$$m' \approx (m - x_l)\cos\delta_l + (n - y_l)\sin\delta_l \tag{6}$$

$$n' \approx -(m - x_l)\sin\delta_l + (n - y_l)\cos\delta_l \tag{7}$$

are the closest integers found to constitute the coordinate of a point in $s_l(m,n)$ at which the intensity is approximately equal to that at the corresponding point in $s_{l+1}(m,n)$, and $\epsilon_T(m',n')$ is the error introduced because of the fixed geometry of the array configuration and the limit to have integers m' and n' for a finite detector array. However, practically speaking, this error should be negligible as compared to the random disturbance from external sources and, therefore, is not considered in the analytical model to be discussed later.

In the transform domain, it can be easily shown that the spatial-domain relationship stated in Eq.(5) becomes

$$\mathcal{S}_{l+1}(m,n) \approx \mathcal{S}_l(\alpha,\beta) e^{i\theta_{m,n,l}} \tag{8}$$

where

$$\alpha \approx m\cos\delta_l + n\left(\frac{M}{N}\right)\sin\delta_l \tag{9}$$

$$\beta \approx n\cos\delta_l - m\left(\frac{N}{M}\right)\sin\delta_l \tag{10}$$

are again the closest integers for the same reason mentioned above, and

$$\theta_{m,n,l} = -2\pi\left(\frac{m}{M}x_l + \frac{n}{N}y_l\right) \tag{11}$$

is the phase introduced to the pixel (m,n) of the transform-domain image due to the translation of coordinate from t_l to t_{l+1}. It is interesting to note that the effects of translational and rotational movement are nicely separated in the transform domain, with the information about image translation contained in the phase and that about image rotation in the coordinate of the transform-domain image. For the cases that no image rotation is involved[3] (i.e., $\delta_l = 0$), we have

$$\mathcal{S}_{l+1}(m,n) = \mathcal{S}_l(m,n) e^{-i2\pi\left(\frac{m}{M}x_l + \frac{n}{N}y_l\right)} \tag{12}$$

as expected.

Estimation of Translation Vector and Rotation Angle

Based on the Gaussian assumption stated in Eq. (1), the maximum likelihood estimator will declare the

[3] This is true in most cases we are interested since the rotation of an imager, if any, is usually a predictable behavior which can be fully compensated.

estimated translation vector $(\hat{x}_l, \hat{y}_l)$ and the rotation angle $\hat{\delta}_l$ if

$$p(\vec{r}_{l+1} \mid \hat{x}_l, \hat{y}_l, \hat{\delta}_l; \vec{s}_l) = \max_{\{x_l, y_l, \delta_l\}} p(\vec{r}_{l+1} \mid x_l, y_l, \delta_l; \vec{s}_l) \quad (13)$$

where, for notational convenience, $\vec{r}_{l+1}$ and $\vec{s}_l$ are the vector representation[4] of the corresponding received image matrix $r_{l+1}(m,n)$ and the reference image matrix $s_l(m,n)$, respectively, and

$$p(\vec{r}_{l+1} \mid x_l, y_l, \delta_l; \vec{s}_l) = \frac{1}{\left(\sqrt{2\pi}\sigma_l\right)^{MN}} e^{-\frac{1}{2\sigma_l^2}\|\vec{r}_{l+1} - \mathbb{L}_{x_l, y_l, \delta_l}\{\vec{s}_l\}\|^2} \quad (14)$$

is the conditional probability density function of $\vec{r}_{l+1}$ given that the translation vector is (x_l, y_l), the rotation angle is δ_l, and the reference at the beginning of this movement is $\vec{s}_l$. Here, $\mathbb{L}_{x,y,\delta}\{\cdot\}$ is defined as a translation/rotation operator which moves the operand by a translation vector (x, y) and rotates it by an angle δ, and the notation $\|\vec{x}\|$ represents a $\mathcal{L}_2$ norm of the vector $\vec{x}$. The maximum likelihood criterion stated in Eq. (13) is equivalent to minimize the exponent in Eq. (14) over all possible (x_l, y_l) and δ_l, rendering the likelihood function to be maximized, after expanding the $\mathcal{L}_2$ norm, as

$$\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} |w_l(m,n)| \cos(\xi_{m,n,l} - \theta_{m,n,l}) \quad (15)$$

where

$$w_l(m,n) \triangleq \mathcal{R}_{l+1}(m,n)\mathcal{S}_l^*(\alpha,\beta) = |w_l(m,n)| e^{i\xi_{m,n,l}} \quad (16)$$

is the pixel-to-pixel product of the transform-domain received image $\mathcal{R}_{l+1}(m,n)$ and the complex conjugate of the transform-domain reference image $\mathcal{S}_l(\alpha,\beta)$, with α and β being given in Eqs. (9) and (10). Note that Eq. (15) can be rewritten as

$$\mathrm{Re}\left\{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} \mathcal{R}_{l+1}(m,n)\mathcal{S}_l^*(\alpha,\beta) e^{-i\theta_{m,n,l}}\right\} \quad (17)$$

where $\mathrm{Re}\{\cdot\}$ represents the real part of a complex quantity. It is clearly indicated that the likelihood function involves the average over all pixels of the pixel-wise multiplied received and reference images in the transform domain, as well as the phase to be estimated.

[4] It is known as the lexicographic form in which the rows (or columns) of a matrix are concatenated sequentially to form a vector.

As suggested by Eq. (13), the translation vector and rotation angle should be jointly estimated by maximizing Eq. (17), which inevitably requires a massive parallel search through every possible coordinate shift and rotation angle since the information regarding the rotational movement is embedded in the coordinate (α, β) of the transform-domain image. However, in reality, this problem can be greatly alleviated somewhat by breaking the one-shot joint estimation into two steps, namely first estimating the translation vector and then searching for the rotation angle with the help from information of the estimated translation. The reason behind this two-step approach is the assumption that the rotational movement, if any, has usually been greatly compensated for and, therefore, is much slower than the transitional movement. A reasonably good convergence is usually achievable after repeating this process several times. In the following, we will concentrate on the algorithm of estimating the translation vector.

By taking the partial derivatives of the likelihood function in Eq. (15) with respect to x_l and y_l and equating them to zero, we have

$$\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} m|w_l(m,n)| \sin(\xi_{m,n,l} - \theta_{m,n,l}) = 0 \quad (18)$$

$$\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} n|w_l(m,n)| \sin(\xi_{m,n,l} - \theta_{m,n,l}) = 0 \quad (19)$$

which form a set of simultaneous nonlinear equations to be solved for the maximum likelihood estimates of x_l and y_l. This set of equations provides the optimal spatial acquisition scheme for rotation-invariant movements.

When the image is close to being acquired, the phase differences, $(\xi_{m,n,l} - \theta_{m,n,l})$, are small and the approximation of $\sin(x) \approx x$ can be applied to Eqs. (18) and (19), rendering a suboptimal linear estimator of which the computation complexity is greatly reduced. The estimated translation vector $(\hat{x}_l, \hat{y}_l)$ obtained from solving this linearized maximum likelihood criterion satisfies

$$\hat{x}_l = \frac{M}{2\pi}\left(\frac{A_{nn}B_m - A_{mn}B_n}{A_{mn}^2 - A_{mm}A_{nn}}\right) \quad (20)$$

$$\hat{y}_l = \frac{N}{2\pi}\left(\frac{A_{mm}B_n - A_{mn}B_m}{A_{mn}^2 - A_{mm}A_{nn}}\right) \quad (21)$$

where

$$A_{mm} = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} m^2|w_l(m,n)|$$

$$A_{mn} = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} mn|w_l(m,n)|$$
$$A_{nn} = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} n^2|w_l(m,n)|$$

and

$$B_m = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} m|w_l(m,n)|\xi_{m,n,l}$$
$$B_n = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} n|w_l(m,n)|\xi_{m,n,l}$$

are coefficients to be calculated based on the information contained in Eq. (16).

Image Tracking Loop

A closed-loop image tracking algorithm motivated by the same maximum likelihood criterion can also be developed for continuous tracking of the translation movement between the imager and target. As discussed in the previous section, to maximize the likelihood function of acquiring an image involves a comparison of the received image against the reference image. However, for image tracking, it is the correlation between the transform-domain received image $\mathcal{R}_{l+1}(m,n)$ and the translated reference image

$$\hat{\mathcal{S}}_{l+1}(m,n) = \mathcal{S}_l(m,n)e^{i\hat{\theta}_{m,n,l}} \tag{22}$$

established from previous estimation to be continuously monitored. The pixel-wise product of $\mathcal{R}_{l+1}(m,n)$ and $\hat{\mathcal{S}}^*_{l+1}(m,n)$ can be expressed by

$$\begin{aligned} C_{l+1}(m,n) &\triangleq \mathcal{R}_{l+1}(m,n)\mathcal{S}^*_l(m,n)e^{-i\hat{\theta}_{m,n,l}} \\ &= |\mathcal{S}_l(m,n)|^2 e^{i\phi_{m,n,l}} \\ &\quad + \mathcal{N}_{l+1}(m,n)\mathcal{S}^*_l(m,n)e^{-i\hat{\theta}_{m,n,l}} \end{aligned} \tag{23}$$

where $\hat{\theta}_{m,n,l}$ is the estimate of $\theta_{m,n,l}$ in Eq. (11) and the estimation error is defined as

$$\begin{aligned} \phi_{m,n,l} &= \theta_{m,n,l} - \hat{\theta}_{m,n,l} \\ &= -2\pi\left[\frac{m}{M}(x_l - \hat{x}_l) + \frac{n}{N}(y_l - \hat{y}_l)\right] \\ &\triangleq -2\pi\left(\frac{m}{M}\Delta_x + \frac{n}{N}\Delta_y\right) \end{aligned} \tag{24}$$

where Δ_x and Δ_y are the associated errors in the estimated coordinate.

Similar to Eq. (17), the real part of Eq. (23) is found to be maximized over the entire detector array, producing the likelihood function

$$\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} \mathrm{Re}\left\{C_{l+1}(m,n)\right\} \tag{25}$$

which is the real part of the correlation between $\mathcal{R}_{l+1}(m,n)$ and $\hat{\mathcal{S}}_{l+1}(m,n)$. It turns out with no surprise that $\hat{x}_l$ and $\hat{y}_l$ can be obtained by solving two simultaneous nonlinear equations formed by setting the partial derivatives of Eq. (25) with respect to $\hat{x}_l$ and $\hat{y}_l$ to be zero. However, in deriving the tracking algorithm to continuously update the estimates, two simultaneous loop feedback signals are formed instead as the partial derivatives of Eq. (25) with respect to Δ_x and Δ_y, rendering

$$\begin{aligned} \varepsilon_x &\triangleq \frac{\partial}{\partial\Delta_x}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\mathrm{Re}\left\{C_{l+1}(m,n)\right\} \\ &= \frac{2\pi}{M}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} m|\mathcal{S}_l(m,n)|^2\sin\left(\phi_{m,n,l}\right) \\ &\quad + \sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\mathcal{N}^{(x)}_{l,eff}(m,n) \end{aligned} \tag{26}$$

$$\begin{aligned} \varepsilon_y &\triangleq \frac{\partial}{\partial\Delta_y}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\mathrm{Re}\left\{C_{l+1}(m,n)\right\} \\ &= \frac{2\pi}{N}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} n|\mathcal{S}_l(m,n)|^2\sin\left(\phi_{m,n,l}\right) \\ &\quad + \sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\mathcal{N}^{(y)}_{l,eff}(m,n) \end{aligned} \tag{27}$$

where

$$\begin{aligned} &\mathcal{N}^{(x)}_{l,eff}(m,n) \\ &\quad = \frac{\partial}{\partial\Delta_x}\mathrm{Re}\left\{\mathcal{N}_{l+1}(m,n)\mathcal{S}^*_l(m,n)e^{-i\hat{\theta}_{m,n,l}}\right\} \end{aligned}$$

$$\begin{aligned} &\mathcal{N}^{(y)}_{l,eff}(m,n) \\ &\quad = \frac{\partial}{\partial\Delta_y}\mathrm{Re}\left\{\mathcal{N}_{l+1}(m,n)\mathcal{S}^*_l(m,n)e^{-i\hat{\theta}_{m,n,l}}\right\} \end{aligned}$$

are the effective noises in the loop operation. Equations (26) and (27) characterize the relationship between the estimate errors, Δ_x and Δ_y, and the loop feedback signals, ε_x and ε_y. However, to solve for Δ_x and Δ_y from these nonlinear equations can be a quite challenging task. With a reasonable assumption valid when the phase error $\phi_{m,n,l}$ remains small during the tracking mode, one can substitute Eq.

(24) for $\sin(\phi_{m,n,l})$ in Eqs. (26) and (27). The resulting simultaneous equations are linear for Δ_x and Δ_y and can be easily solved, yielding

$$\Delta_x = \frac{C_n \mathbb{E}[\varepsilon_x] - C_{mn}\mathbb{E}[\varepsilon_y]}{C_m C_n - C_{mn}^2} \tag{28}$$

$$\Delta_y = \frac{C_{mn} \mathbb{E}[\varepsilon_x] - C_m \mathbb{E}[\varepsilon_y]}{C_{mn}^2 - C_m C_n} \tag{29}$$

where $\mathbb{E}[\cdot]$ denotes the statistical expectation and

$$C_m \triangleq \frac{4\pi^2}{M^2} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} m^2 |\mathcal{S}_l(m,n)|^2$$

$$C_n \triangleq \frac{4\pi^2}{N^2} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} n^2 |\mathcal{S}_l(m,n)|^2$$

$$C_{mn} \triangleq \frac{4\pi^2}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} mn |\mathcal{S}_l(m,n)|^2$$

are coefficients that can be calculated from the transform-domain reference image of the previous iteration at t_l.

An extended-image tracking loop structure can be realized based upon the above derivation and is depicted in Fig. 2. The transform-domain received image $\{\mathcal{R}_{l+1}(m,n)\}$ is first multiplied pixel-wise with a properly translated transform-domain reference image established according to the estimate $(\hat{x}_l, \hat{y}_l)$ from the previous iteration at t_l. After being averaged over the entire detector array, the correlation result is used to compute the loop feedback signals, ε_x and ε_y. It is important to note that, in real implementation, these two loop feedback signals are calculated from $C_{l+1}(m,n)$ as the following weighted sums

$$\varepsilon_x = \frac{2\pi}{M} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} m \mathrm{Im}\{C_{l+1}(m,n)\} \tag{30}$$

$$\varepsilon_y = \frac{2\pi}{N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} n \mathrm{Im}\{C_{l+1}(m,n)\} \tag{31}$$

where $\mathrm{Im}\{\cdot\}$ represents the imaginary part of a complex quantity since Δ_x and Δ_y are not readily accessible for the partial derivatives given in Eq. (26) and (27). The subsequent calculation of Δ_x and Δ_y from ε_x and ε_y is straightforward as indicated in Eqs. (28) and (29), except that the statistical averages are replaced by time averages performed through low-pass filtering. The calculated Δ_x and Δ_y are then used to update the movement estimates through an accumulator, such that

$$\hat{x}_{l+1} = \hat{x}_l + \Delta_x \tag{32}$$

$$\hat{y}_{l+1} = \hat{y}_l + \Delta_y \tag{33}$$

The updated accumulator contents will be used to calculate the estimate $\hat{\theta}_{m,n,l+1}$ and prepare the translated reference image for the next loop iteration at t_{l+1}.

NUMERICAL RESULTS

The proposed transform-domain correlation-type spatial acquisition and tracking technique can be applied to many areas, especially for those with asymmetric target profiles or very low signal-to-noise ratios where the conventional spatial-domain centroid-type algorithms fail. For example, in deep-space laser communications, it can be used in both the optical transceiver onboard the spacecraft and the Earth-based optical terminal. In the former case, this scheme enables the use of a Sun-lit Earth image (or an image of Earth and Moon altogether) as the beacon signal for spacecraft to determine the location of the ground terminal. In the latter case, a priorly known intensity profile from a distant point-source (e.g., a star) is used by the Earth-based optical terminal as the reference image for simultaneous beam tracking and telemetry data detection on a high-speed detector array.

The numerical results presented as follows are intended to demonstrate its capability to achieve sub-pixel accuracy in estimating and tracking rotation-invariant movements in high disturbance environments. No rotational effect is included in the simulations.

Acquisition of Earth Beacon

The spatial acquisition algorithm has been simulated for different Sun-lit Earth images shown in Fig. 3. The 64×64 reference image is compressed from a 256×256 original taken by the Galileo spacecraft during its mission to Jupiter. The received image is assumed to be detected by a 16×16 CCD array and corrupted by additive white Gaussian disturbances such that the average signal-to-noise ratio (SNR) defined as

$$\left(\frac{S}{N}\right) \triangleq \frac{1}{\sqrt{MN}} \frac{\| \vec{s}_l \|}{\sigma_l} \tag{34}$$

is unity in this simulation.

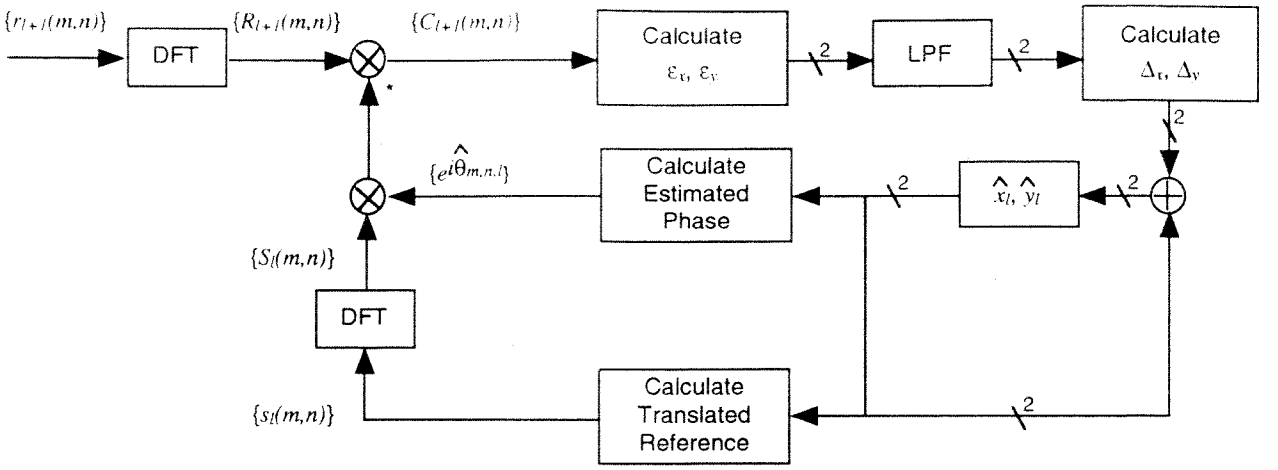
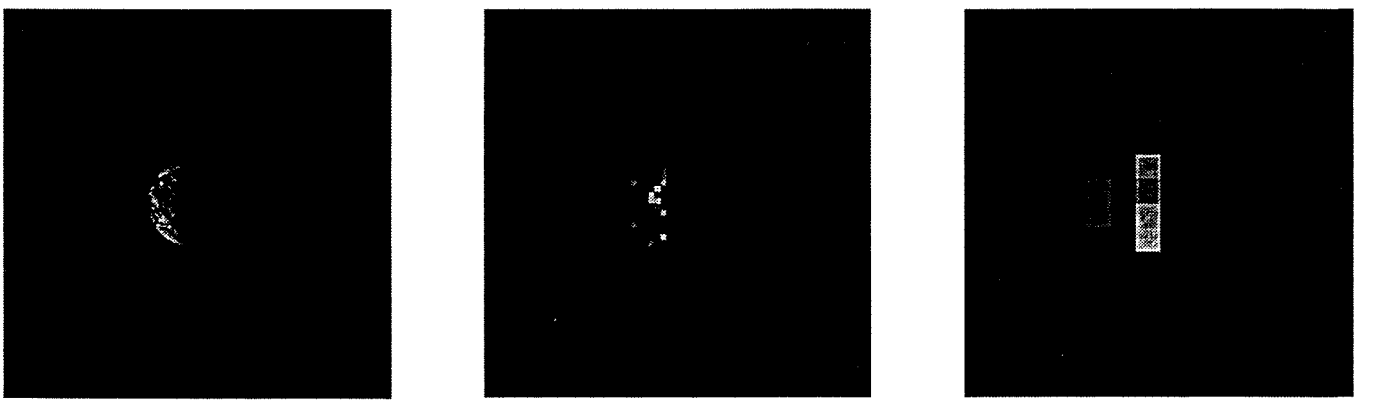


Figure 2: The Extended-Image Tracking Loop.



(a) (b) (c)

Figure 3: The Extended-Source Images: (a) original image of size 256 × 256, (b) reference image of size 64 × 64, and (c) received image by a detector array of size 16 × 16.

Computation Index		Estimated Coordinate		Actual Coordinate	
M_c	N_c	Rows	Columns	Rows	Columns
4	4	3.75	-3.14	3.625	-2.82
6	6	3.55	-2.84		
8	8	3.55	-2.81		

Table 1: Estimated Ground Station Coordinate during Acquisition

During acquisition, the reference image is scaled down to a 16×16 image to match the size of the received image. The acquisition process as suggested by Eqs. (18) and (19) requires $M \times N$ terms of summation for each of the simultaneous equations. The actual size of computation can be reduced by first forming partial sums of these equations and then gradually including more terms in those partial sums until estimated results converge. Table 1 shows the estimated coordinate as the size of computation, specified by M_c and N_c, increases. The estimates are refined to (3.55,-2.81) relative to the 16×16 CCD array, which translates to a sub-pixel accuracy of less than 2 and 1 percent in the respective direction of the original reference image.

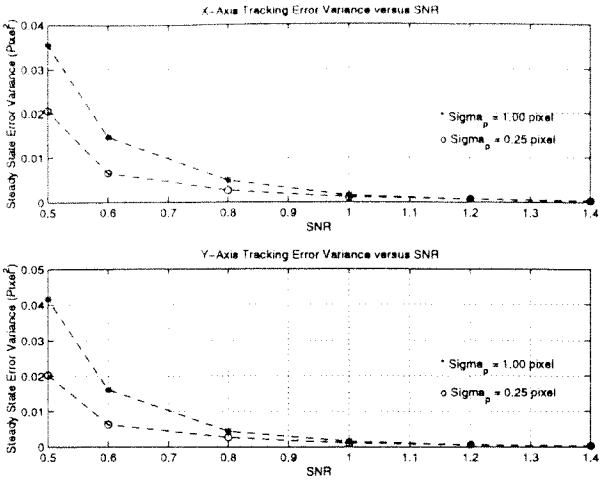


Figure 4: The Error Variance for Image Tracking.

Laser Beam Tracking

The image tracking algorithm has been simulated for the optical downlink. In this simulation, the detected intensity profile of the laser spot is assumed to be a two-dimensional Gaussian pulse with its spread specified by the standard deviation, denoted as σ_p, of the Gaussian probability density function. It is further assumed that a 4×4 detector array is used to capture the laser spot image. In the simulation, the laser spot is initially offset by one pixel in both x- and y-direction on the detector array, rendering only about two-thirds of the laser power being collected by the array when $\sigma_p = 1.0$ pixel versus over 99% of power when $\sigma_p = 0.25$ pixel.

Figure 4 shows the tracking error variance versus various SNRs defined in Eq. (34) for different σ_p. It appears that tracking of a less-concentrated beam (e.g., $\sigma_p = 1.0$ pixel versus $\sigma_p = 0.25$ pixel) suffers no significantly more degradation except in the extremely low SNR region, which appears to be a key advantage over the spatial-domain centroid-type tracking algorithms. The numerical results indicate that an exponentially decay of error variance as SNR increases, with the standard deviation of the tracking error being 3.32×10^{-2} pixel for $\sigma_p = 0.25$ pixel and 3.87×10^{-2} pixel for $\sigma_p = 1$ pixel when SNR equals to unity.

CONCLUSION

Accurate and stable optical pointing to a moving target has been an essential function in many laser-based applications. This paper describes in details a generic maximum-likelihood-based transform-domain correlation-type extended-image spatial acquisition and tracking scheme which has been demonstrated through numerical simulations to be able to achieve sub-pixel accuracy in estimating and tracking relative translation movement between the target and the imager in high disturbance environments. With an additional rotation angle estimation assisted by the information from the translation vector estimation to compensate for possible image rotation, this scheme can provide a cost-effective way to enhance pointing accuracy without putting too much extra burden on the imager's pointing control mechanism.

Although the optimal solution (in the sense of maximum likelihood) to the spatial acquisition and tracking has been analytically established in this study, there are still areas remaining for further investigation. For example, the need for adaptation of the reference profile to compensate for line-of-sight optical turbulence is an important issue to be resolved before successfully applying this technique in a high dynamic situation. Furthermore, the assumption of independent white Gaussian disturbance in each pixel does not address many different types of disturbances. For example, sub-pixel scanning technique for resolution enhancement and various atmospheric effects can create spatially correlated disturbances which will be interesting topics for future studies. More realistic simulation and demonstration focusing on small-sized focal plane array and real laser beam profile subject to optical turbulence are required to determine the relationship between system performance and important design parameters such

as required detector sensitivity and minimum loop update rate, etc.

References

[1] R. M. Gagliardi and S. Karp, *Optical Communications*, John Wiley & Sons, New York, 2 ed., 1995.

[2] F. G. Smith, ed., *Atmospheric Propagation of Radiation*, vol. 2 of *The Infrared and Electro-Optical Systems Handbook*, Infrared Information Analysis Center, Ann Arbor, Michigan and SPIE Optical Engineering Press, Bellingham, Washington, 1993.

[3] C. D. Kuglin and D. C. Hines, "The phase correlation image alignment method," *Proc. Int. Conf. on Cybernetics and Society* , pp. 163–165, 1975.

[4] J. J. Pearson, D. C. Hines, S. Golosman, and C. D. Kuglin, "Video-rate image correlation processor," in *Proc. SPIE*, vol. 119, Application of Digital Image Processing, IOCC 1977, pp. 197–205, 1977.

[5] E. D. Castro and C. Morandi, "Registration of translated and rotated images using finite fourier transforms," *IEEE trans. on Pattern Analysis and Machine Intelligence* **PAMI-9**, pp. 700–703, September 1987.